Detectores de Textos Gerados por Inteligência Artificial na Avaliação da Aprendizagem: Uma Análise Comparativa em Contextos Educacionais

João Vitor da Silva Teixeira[1], Daniel dos Anjos Costa[2]

[1] Instituto Federal da Bahia (IFBA) – Ubaitaba – BA – Brasil

[2] Instituto Federal da Bahia (IFBA) – Santo Amaro – BA – Brasil

{joaovitor771809@gmail.com e daniel.anjos@ifba.edu.br}

Abstract. *This article addresses the challenges of learning assessment in the era of Generative Artificial Intelligences (GAIs), such as ChatGPT and Gemini, focusing on the effectiveness and limitations of AI text detection tools. The methodology employed a comparative and exploratory analysis of seven detectors (Justdone, ZeroGPT, Detecting-AI, Smodin.io, Copyleaks, Quillbot, and Rephrase) across five textual categories: original human writing, human text altered by AI, content entirely produced by AI, a technical report in another language, and human speech transcription. The results revealed significant inconsistencies in the accuracy of the tools, reinforcing hybrid approaches for detecting AI-generated texts.*

Resumo. *Este artigo aborda os desafios da avaliação de aprendizado na era das Inteligências Artificiais Generativas (IAGs), como ChatGPT e Gemini, focando na eficácia e limitações de ferramentas de detecção de IA em textos. A metodologia empregou uma análise comparativa e exploratória de sete detectores (Justdone, ZeroGPT, Detecting-AI, Smodin.io, Copyleaks, Quillbot e Rephrase) em cinco categorias textuais: escrita humana original, texto humano alterado por IA, conteúdo totalmente gerado por IA, relatório técnico em outro idioma e transcrição de fala humana. Os resultados revelaram inconsistências significativas na acurácia das ferramentas reforçando abordagens híbridas para detecção de textos gerados por IA.*

1. Introdução

As Inteligências Artificiais Generativas (IAGs), como *ChatGPT*, *Gemini* e *Bard*, têm transformado os processos de produção textual e provocado novos desafios para a avaliação da aprendizagem. Em contextos educacionais, torna-se cada vez mais difícil distinguir textos produzidos integralmente por estudantes daqueles gerados ou modificados por sistemas de IA. Como aponta Correia (2023), conteúdos produzidos por inteligência artificial já circulam em ambientes acadêmicos e científicos, muitas vezes sem identificação imediata de sua origem.

Nesse cenário, emerge a seguinte questão de pesquisa: como avaliar, de forma justa e confiável, a aprendizagem dos estudantes diante de textos que podem ser parcial ou totalmente elaborados por IA? A problemática amplia discussões já existentes sobre autoria acadêmica, uma vez que, conforme Krokoscz (2015), a verificação de autoria e originalidade textual já representava um desafio antes mesmo da popularização das IAGs.

Diante disso, este artigo analisa o desempenho de sete ferramentas de detecção de textos gerados por IA — *Justdone*, *ZeroGPT*, *Detecting-AI*, *Smodin.io*, *Copyleaks*, *Quillbot* e *Rephrase* — aplicadas a diferentes categorias textuais: textos humanos, textos modificados por IA, textos totalmente gerados por IA, relatório técnico em outro idioma

e transcrição de fala humana. O objetivo é verificar as potencialidades e limitações desses detectores, especialmente quanto à ocorrência de falsos positivos e falsos negativos em situações próximas ao contexto educacional.

A relevância do estudo está na necessidade de repensar práticas avaliativas em um cenário de ampla adoção de tecnologias baseadas em IA. Assim, busca-se contribuir para o debate sobre o uso responsável de detectores automáticos, defendendo sua aplicação como apoio à avaliação docente, e não como critério definitivo de julgamento da autoria textual.

2. Referencial Teórico

Esta seção estabelece o alicerce teórico deste estudo. Nela, faremos uma revisão da literatura relevante, explorando os principais conceitos e ideias que servem de base para a pesquisa. O objetivo é fornecer o suporte conceitual necessário para que a análise dos dados sejam claras e bem fundamentadas.

2.1 Instrumentos para identificar inteligência artificial em escritos

Com o aumento da utilização de inteligências artificiais generativas (IAGs) na criação de textos, há uma demanda por ferramentas talentosas para conseguir distinguir textos humanos de textos que são gerados por IA. Esse dilema está ocasionando preocupação e vem ganhando realces em ambientes acadêmicos, artigos e editoriais, instigando o avanço de métodos e ferramentas típicas para essa finalidade.

Os softwares de IA são sistemas que examinam padrões de linguagem, configurações gramaticais, medições estatísticas e estilísticas para deduzir se um texto foi criado por uma pessoa ou por uma ferramenta de linguagem automatizada. Segundo Gao et al. (2023), diversas ferramentas desse segmento se fundamentam em Modelos de Classificação Binária desenvolvidos para identificar padrões comuns em textos elaborados por LLMs (Modelos de Linguagem de Grande Escala). Tais padrões incluem repetições, probabilidades de palavras e fluidez acentuada.

Entre as ferramentas mais citadas destaca-se o *GPTZero*, que foi desenvolvido por Edward Tian, que busca avaliar a "perplexidade" e a "*burstiness*" da escrita textual — ações que avaliam a consistência e a diversidade das palavras. Documentos que apresentam baixa complexidade e mínima mudança são frequentemente identificados como produzidos por inteligência artificial.

Outros detectores amplamente conhecidos incluem o *OpenAI AI Text Classifier*, o *Turnitin AI Detection* e o *ZeroGPT*, cada um utilizando algoritmos diferentes, mas com a meta compartilhada de identificar padrões sintéticos. Apesar disso, estudos apontam entraves significativos na eficácia desses detectores. Liang et al. (2023) demonstram que muitos desses sistemas têm índices consideráveis de falsos positivos, ou seja, identificam erroneamente textos escritos por humanos como se fossem gerados por máquinas, principalmente em conteúdos de autores não nativos. Além do mais, instrumentos de redigitação automática e uso planejado de sinônimos são capazes de driblar os detectores, ocasionando uma desconfiança em sua confiabilidade em âmbitos críticos, como apreciações escolares ou dependências acadêmicas.

Por essa razão, especialistas sugerem métodos que sejam combinados e integrados, que ofereçam uma mescla de avaliação automática com discernimento humano, além de hábitos pedagógicos lúcidos sobre a utilização de IA. Segundo Silva e Gomes (2023), educadores conseguem implementar ferramentas de detecção não como instrumento punitivos, todavia como forma de diálogo e observação sobre criação, comunicação e moral na elaboração de textos.

Com o progresso constante da inteligência artificial, os métodos de identificação devem se desenvolver ao mesmo tempo, incluindo modelos que sejam mais fortes, flexíveis e sensíveis ao contexto. O problema se apresenta não apenas do ponto de vista técnico, mas também epistemológico: como reconhecer o "estilo humano" em um tempo em que as máquinas conseguem imitar quase de forma impecável.

3. Trabalhos Relacionados

Pesquisas recentes têm examinado a confiabilidade de ferramentas destinadas à identificação de textos produzidos por Inteligência Artificial, especialmente considerando o avanço dos modelos de linguagem de grande escala. Gao et al. (2023) analisam métodos de detecção fundamentados em padrões linguísticos e preservação de conteúdo, enquanto Liang et al. (2023) destacam limitações relevantes desses detectores, como a incidência de falsos positivos e falsos negativos em distintos contextos de escrita.

No contexto educacional, tais limitações tornam-se ainda mais relevantes, uma vez que a classificação incorreta de um texto pode comprometer a avaliação da aprendizagem e resultar em interpretações injustas acerca da autoria estudantil. Portanto, os detectores automáticos devem ser considerados instrumentos auxiliares, e não mecanismos conclusivos de verificação de autoria.

Em contraste com estudos focados exclusivamente no desempenho técnico dos detectores, este trabalho investiga o comportamento dessas ferramentas em categorias textuais alinhadas ao contexto educacional, incluindo textos humanos, textos modificados por IA, textos integralmente gerados por IA, relatórios técnicos em outros idiomas e transcrições de fala humana. Dessa forma, a principal contribuição deste artigo consiste em evidenciar as inconsistências dessas ferramentas e discutir seus impactos no processo de avaliação da aprendizagem.

4. Metodologia

Este trabalho caracteriza-se por sua natureza aplicada, adotando uma abordagem qualitativa e exploratória. A investigação concentrou-se na análise comparativa de ferramentas de detecção de Inteligência Artificial (IA) em textos, com o propósito central de avaliar o desempenho e a eficácia na identificação de trechos gerados por Inteligências Artificiais Generativas (IAGs), como *ChatGPT* e *Gemini*. Sete ferramentas foram analisadas neste trabalho, a saber: *ZeroGPT*, *Copyleaks*, *Justdone*, *Quillbot*, *Smodin.io*, *Detecting-AI* e *Rephrase*.

Para tanto, a pesquisa foi conduzida com uma amostragem composta por cinco categorias distintas de conteúdo textual, cada uma representando variados níveis de interferência de Inteligência Artificial:

1. Trecho de artigo acadêmico original (escrito por humano);
2. Trecho humano alterado com a contribuição de Inteligência Artificial;
3. Texto totalmente produzido por Inteligência Artificial;
4. Trecho de relatório técnico em outro Idioma;
5. Texto transcrito de áudio para análise do comportamento das ferramentas.

5. Resultados e Discussão

Esta seção apresenta os resultados obtidos a partir dos testes realizados com as ferramentas de detecção de Inteligência Artificial (IA). Serão detalhadas as performances observadas e as variações significativas entre as análises, mesmo quando aplicadas ao mesmo conteúdo textual. Os achados serão discutidos em etapas distintas, oferecendo uma compreensão aprofundada da eficácia de cada ferramenta.

5.1 Avaliação das ferramentas com texto de Autoria Humana

Neste primeiro teste, foi utilizado um trecho do artigo **"**Textos Gerados por Inteligência Artificial e suas Implicações no EAD",O Quadro 1 apresenta o fragmento do texto utilizado na análise das ferramentas.

Excerto 1 - Trecho do texto acadêmico original. Fonte: Almeida, (2023)

> "*Com o lançamento de revolucionários programas geradores de texto, de altíssima qualidade, baseados em inteligência artificial, diversas implicações começaram a surgir, afetando diretamente e, de forma significativa, a educação ao redor do mundo.*
> *Conforme afirmado por Bervian, Cervo e Silva (2006), a pesquisa, o acesso ao conhecimento e a produção de novas informações evoluem a passos largos. Carlini (2022) exemplifica que o avanço tecnológico é tão rápido que, há 10 anos, muitas das coisas que usamos hoje e nos parecem nativas, não existiam, citando, como exemplo, aplicativos de mensagens instantâneas, atualmente utilizados pela grande maioria das pessoas que possuem um aparelho celular. Cita, ainda, o uso da inteligência artificial, que auxilia o ser humano em diversas áreas do conhecimento. O autor, inclusive, faz uma analogia de que, atualmente, a tecnologia está tão inserida nos nossos dias que parece fazer parte do nosso DNA.*
> *Delimitando a evolução tecnológica na área de inteligência artificial, com foco no seu uso em programas geradores de texto, se, por um lado, o avanço da tecnologia é visto de forma positiva, sendo muito benéfico aos que a utilizam com ética para explorar novos horizontes, pesquisar e aprender de modo a desenvolver e compartilhar trabalhos de pesquisa científica com excelência, de outro, é visto de forma negativa, pois, infelizmente, uma parcela a utiliza com a intenção de gerar conteúdo de forma rápida e sem grandes esforços, sem verificar a qualidade e procedência do conteúdo gerado, o que, por consequência, fere os princípios da ética no mundo acadêmico e pode, inclusive, incorrer em plágio.*"

O quadro 01 apresenta o resultado da análise em formato gráfico. Conforme resultados, a ferramenta *Justdone* indicou uma predominância significativa de conteúdo produzido por IA, classificando o texto como 85% "Feito por IA" e apenas 15% "Escrito por humano". Esse resultado representa um falso positivo substancial, dado que o texto é de autoria humana.

Conforme sintetizado no Quadro 1, *ZeroGPT*, *Detecting-AI*, *Smodin.io*, *Copyleaks* e *Rephrase* classificaram corretamente o texto como 100% escrito por humano, demonstrando maior consistência na identificação de autoria humana no cenário analisado. A *Quillbot* apresentou resultado intermediário, indicando 69% de autoria humana e 31% de IA, o que caracteriza um falso positivo parcial. Já a *Justdone* apresentou o resultado mais crítico, ao classificar 85% do texto como produzido por IA, mesmo se tratando de um conteúdo integralmente humano.

Esses achados evidenciam que alguns detectores podem atribuir autoria artificial a textos humanos, reforçando a necessidade de cautela no uso dessas ferramentas como critério isolado em processos de avaliação da aprendizagem.

Quadro 1 – Síntese da avaliação das ferramentas com texto de autoria humana.

Ferramenta	Feito por IA (%)	Escrito por humano (%)	Interpretação
Justdone	85%	15%	Falso positivo elevado
ZeroGPT	0%	100%	Classificação correta
Detecting-AI	0%	100%	Classificação correta
Smodin.io	0%	100%	Classificação correta
Copyleaks	0%	100%	Classificação correta
Quillbot	31%	69%	Indício de falso positivo parcial
Rephrase	0%	100%	Classificação correta

5.2 Avaliação com texto de autoria humana com modificações feitas por IA

Nessa avaliação, foi empregado um conteúdo de natureza instrutiva disponível na plataforma Toda Matéria e de autoria de Thiago Souza (educador nas áreas de Sociologia, Filosofia e História) com modificações feitas pelo *ChatGPT*. O excerto 2 apresenta o trecho original do texto e o excerto 3 exibe o trecho modificado pelo *ChatGPT*.

Excerto 2 - Trecho do texto original. Fonte: Pará (2024).

> *"A história do Brasil começou com a ocupação dos seres humanos há cerca de 12-20 mil anos. No século XVI, os portugueses começaram a colonizar estas terras e transferiram africanos para serem mão de obra escrava nos engenhos que construíram aqui. Por sua vez, estes trabalhadores forçados trariam novos alimentos e animais que mudariam a história dos povos originários para sempre."*

Excerto 3 - Trecho modificado com auxílio do *ChatGPT*.

> *"A história do Brasil tem início com a ocupação humana, ocorrida há aproximadamente 12 a 20 mil anos. No século XVI, os portugueses iniciaram o processo de colonização dessas terras, trazendo africanos para atuarem como mão de obra escravizada nos engenhos que estabeleceram. Esses trabalhadores forçados, por sua vez, introduziram novos alimentos e animais, transformando de maneira definitiva a história e a cultura dos povos originários."*

O quadro 2 ilustra os resultados obtidos neste teste. A ferramenta *Justdone* demonstrou um claro reconhecimento do conteúdo como sendo gerado por inteligência artificial, classificando 85% como "Feito por IA" e apenas 15% como "Escrito por humano". Esse resultado indica que, na percepção do *Justdone*, as modificações introduzidas pelo *ChatGPT* mantiveram traços significativos de geração por IA no texto.

As ferramentas *ZeroGPT*, *Copyleaks*, *Quillbot* e *Rephrase* indicaram de forma consistente uma alta e uniforme presença de conteúdo criado artificialmente, classificando-o como 100% "Feito por IA". A unanimidade e objetividade desses resultados indicam que, para essas ferramentas, o texto, mesmo após as alterações, reteve características que o categorizam de forma absoluta como produzido por inteligência artificial.

A ferramenta *Detecting-AI* exibiu um padrão de detecção distinto, reconhecendo 37% como "Feito por IA" e 63% como "Escrito por Humano". Essa divergência em relação à tendência predominante sugere que a *Detecting-AI* emprega um algoritmo de

avaliação que tem uma maior propensão a identificar a autoria humana ou a desconsiderar certos padrões de IA que outras ferramentas conseguem detectar.

Quadro 2 – Síntese da avaliação das ferramentas com texto de autoria humana com modificações feitas por IA.

Ferramenta	Feito por IA (%)	Escrito por humano (%)	Interpretação
Justdone	85%	15%	Forte indício de interferência de IA
ZeroGPT	100%	0%	Identificação integral de IA
Detecting-AI	37%	63%	Detecção parcial da interferência de IA
Smodin.io	0%	100%	Não identificou interferência de IA
Copyleaks	100%	0%	Identificação integral de IA
Quillbot	100%	0%	Identificação integral de IA
Rephrase	100%	0%	Identificação integral de IA

Por fim, a ferramenta *Smodin.io* se destacou por não identificar qualquer sinal de Inteligência Artificial, categorizando o texto integralmente como "Escrito por Humano". Este resultado é particularmente relevante, pois sugere que, na análise do *Smodin.io*, as modificações realizadas pelo *ChatGPT* não foram suficientes para que o software detectasse a presença de Inteligência Artificial no texto

De modo geral, este caso demonstrou uma maior sensibilidade na detecção por parte das ferramentas — especialmente *ZeroGPT*, *Copyleaks*, *Quillbot* e *Rephrase*, que apresentaram elevada capacidade em identificar a intervenção artificial no texto.

5.3. Avaliação das ferramentas com texto totalmente gerado por IA

Neste estágio, o objetivo foi analisar a performance das ferramentas ao serem aplicadas a um texto inteiramente criado pela inteligência artificial *ChatGPT*. O texto está sendo exibido no excerto 4 abaixo e teve o tema “Segunda Guerra Mundial”.

Excerto 4 - Trecho do texto totalmente gerado por IA. Fonte: *ChatGPT*.

> *“A Segunda Guerra Mundial foi um dos maiores e mais devastadores conflitos da história, ocorrido entre 1939 e 1945. Envolveu a maioria das nações do mundo, divididas em dois grandes blocos: os Aliados (como Estados Unidos, União Soviética e Reino Unido) e o Eixo (liderado por Alemanha, Itália e Japão). O conflito teve início com a invasão da Polônia pela Alemanha nazista, sob o comando de Adolf Hitler. Ao longo dos anos, milhões de pessoas morreram, incluindo civis vítimas do Holocausto e de bombardeios. A guerra terminou com a rendição da Alemanha em maio de 1945 e do Japão em setembro do mesmo ano, após os bombardeios atômicos de Hiroshima e Nagasaki. As consequências da guerra foram profundas, redefinindo fronteiras, estruturas políticas e a ordem mundial.”*

Conforme apresentado no Quadro 3, *ZeroGPT*, *Copyleaks* e *Quillbot* classificaram corretamente o texto como 100% feito por IA, enquanto a *Justdone* também apresentou desempenho satisfatório, indicando 81% de autoria artificial. Em contrapartida, a *Detecting-AI* identificou apenas 15% do conteúdo como IA, configurando falso negativo expressivo, já que o texto foi integralmente gerado pelo *ChatGPT*. Esses resultados evidenciam diferenças significativas de sensibilidade entre os detectores na identificação de textos totalmente artificiais.

No Quadro 3, destaca-se a inconsistência das ferramentas Smodin.io e Rephrase, que classificaram como 100% escrito por humano um texto integralmente gerado pelo *ChatGPT*. Esse resultado configura falso negativo total e evidencia limitações relevantes desses detectores na identificação de autoria artificial.

Quadro 3 – Síntese da avaliação das ferramentas com texto gerado pelo ChatGPT.

Ferramenta	Feito por IA (%)	Escrito por humano (%)	Interpretação
Justdone	81%	19%	Forte indício de autoria por IA
ZeroGPT	100%	0%	Identificação integral de IA
Detecting-AI	15%	85%	Detecção insuficiente de IA
Smodin.io	0%	100%	Não identificou autoria por IA
Copyleaks	100%	0%	Identificação integral de IA
Quillbot	100%	0%	Identificação integral de IA
Rephrase	0%	100%	Não identificou autoria por IA

Assim, mesmo em um cenário de geração integral por IA, algumas ferramentas não reconheceram a procedência do texto, reforçando que detectores automáticos devem ser utilizados apenas como apoio, e não como critério definitivo em processos avaliativos.

5.4 Avaliação das ferramentas com texto de um relatório técnico em outro Idioma

Nesta etapa, foi analisado um trecho em inglês do relatório técnico *Impacts of the Digital Transformation on the Environment and Sustainability*, de Liu et al. (2020), com o objetivo de verificar o comportamento dos detectores diante de um texto técnico, formal e em outro idioma.

O relatório discute os impactos da transformação digital sobre o meio ambiente e a sustentabilidade, servindo como subsídio para debates e diretrizes ambientais no contexto europeu. O trecho selecionado para análise é apresentado no Quadro 5.

Excerto 5 - Trecho do texto de um relatório técnico em outro Idioma. Fonte: Liu, et al. (2020)

> "*Mobilising industry for a clean and circular economy: The (non-energy) environmental opportunities arising from digitalisation can play an important role in relation circular economy, especially with respect to tackling the issue of electronic waste. Most importantly, the technological advancement plays a role in better collection and subsequent recycling of electronic waste and the reuse of the materials used. For example, the advancement in technology, namely the introduction of smartphones and mobile applications encourages consumers to recycle e-waste at official locations in return for financial incentive.*"

Conforme apresentado no Quadro 4, *Justdone* e *Detecting-AI* apresentaram falsos positivos elevados ao classificarem, respectivamente, 80% e 99,7% de um relatório técnico humano como conteúdo produzido por IA. Por outro lado, *ZeroGPT, Smodin.io, Copyleaks* e *Quillbot* identificaram corretamente o trecho como 100% escrito por humano, enquanto a Rephrase indicou 97% de autoria humana. Esses resultados evidenciam que textos técnicos, formais ou em outro idioma podem gerar interpretações divergentes entre detectores, reforçando a necessidade de cautela em seu uso acadêmico e avaliativo.

Por último, a ferramenta *Rephrase* apontou que 97% do texto foi "Escrito por humano" e apenas 3% foi "Feito por IA". Diferentemente da *Detecting-AI*, este resultado sugere uma presença mínima, mas detectável, de IA, com o conteúdo ainda sendo majoritariamente de origem humana. Isso pode indicar a capacidade da ferramenta de perceber sutis diferenças de estilo ou formatos que ela relaciona a modelos de linguagem.

Quadro 4 – Síntese da avaliação das ferramentas com texto de relatório técnico em outro idioma.

Ferramenta	Feito por IA (%)	Escrito por humano (%)	Interpretação
Justdone	80%	20%	Falso positivo elevado
ZeroGPT	0%	100%	Classificação correta
Detecting-AI	99,7%	0,3%	Falso positivo elevado
Smodin.io	0%	100%	Classificação correta
Copyleaks	0%	100%	Classificação correta
Quillbot	0%	100%	Classificação correta
Rephrase	3%	97%	Classificação correta com indício residual de IA

5.5 Avaliação das ferramentas com transcrição de Áudio de fala humana

Essa etapa do trabalho teve como objetivo avaliar as ferramentas de identificação de IA em um contexto de autoria claramente humana. Para esse propósito, uma gravação de áudio foi feita reunindo informações variadas sobre a cidade de Ubaitaba-BA (incluindo dados como a data, ações da prefeitura, clima, número de habitantes e a fama da cidade no esporte de canoagem) e, em seguida, foi transcrita de forma literal.

Esse procedimento assegura que a origem do material é exclusivamente humana, funcionando como um parâmetro para medir a eficácia das ferramentas em reconhecer textos que não são gerados por IA. O Quadro 6 apresenta o texto da transcrição do áudio e a Figura 5 exibe os resultados.

Excerto 6 - Trecho da transcrição de Áudio de Fala Humana.

> *"Ubaitaba dia 11.06.2025 estamos aqui presente na prefeitura municipal de Ubaitaba, realizando tarefas e atividades diárias em um dia ensolarado com previsão de fracas chuvas, para região, mas estamos em exercício em pleno vapor até as 17h, ubaitaba uma região com um pouco mais de 18 mil habitantes, cidade das canoas, cidade da canoagem se torna presente não só no Brasil, mas no mundo"*

A ferramenta *Justdone* identificou que 83% do conteúdo foi classificado como "Feito por IA" e 17% como "Escrito por Humano". Esse resultado é particularmente interessante, pois, em um texto de autoria inequivocamente humana (proveniente de uma transcrição de fala), a *Justdone* ainda atribuiu a maior parte da autoria à IA. Isso indica que seus algoritmos podem captar padrões de linguagem que, mesmo em uma fala transcrita, são equivocadamente associados à produção artificial.

Quadro 5 – Síntese da avaliação das ferramentas com transcrição de áudio de fala humana.

Ferramenta	Feito por IA (%)	Escrito por humano (%)	Interpretação
Justdone	83%	17%	Falso positivo elevado
ZeroGPT	0%	100%	Classificação correta
Detecting-AI	0%	100%	Classificação correta
Smodin.io	0%	100%	Classificação correta
Copyleaks	0%	100%	Classificação correta
Quillbot	0%	100%	Classificação correta
Rephrase	0%	100%	Classificação correta

As ferramentas *ZeroGPT*, *Detecting-AI*, *Smodin.io*, *Copyleaks* e *Quillbot* concordam em classificar o trecho como totalmente elaborado por um ser humano, sem

qualquer vestígio de inteligência artificial (100% "Escrito por Humano"). A unanimidade entre essas cinco plataformas é uma evidência convincente de que reconheceram corretamente a autoria humana do texto. Este desempenho é fundamental para a credibilidade dessas ferramentas, demonstrando que conseguem distinguir com precisão o conteúdo genuíno neste contexto particular. Por fim, a ferramenta Rephrase sinalizou 98% de autoria humana e 2% de autoria por inteligência artificial.

5.6 Síntese dos resultados

A partir dos cinco cenários analisados, foi elaborada uma síntese estatística descritiva considerando a média de indicação como IA, o desvio padrão e a taxa de erro aproximada por ferramenta. Essa sistematização permite observar padrões de comportamento, variação e tendência de erro entre os detectores avaliados.

Quadro 6 – Síntese estatística descritiva por ferramenta.

Ferramenta	Média de Indicação como IA	Desvio padrão	Taxa de erro aproximada	Síntese interpretativa
Justdone	82,8%	2,3	60%	Tendência elevada a classificar textos como IA
ZeroGPT	40,0%	54,8	0%	Bom desempenho geral, com alta variação conforme o tipo textual
Detecting-AI	30,3%	41,6	40%	Comportamento instável, com falso positivo e falso negativo
Smodin.io	0,0%	0,0	40%	Tendência a classificar os textos como humanos
Copyleaks	40,0%	54,8	0%	Bom desempenho geral nas categorias analisadas
Quillbot	46,2%	50,7	20%	Bom desempenho, com falso positivo parcial em texto humano
Rephrase	20,6%	44,4	20%	Falha na identificação de texto integralmente gerado por IA

De modo geral, *ZeroGPT* e *Copyleaks* apresentaram os resultados mais consistentes no conjunto analisado, enquanto *Justdone*, *Detecting-AI*, *Smodin.io* e *Rephrase* evidenciaram limitações relevantes em determinados cenários. Esses achados reforçam que detectores automáticos podem auxiliar a avaliação, mas não devem ser utilizados como critério único para comprovar autoria textual.

6. Conclusão

Este estudo permitiu entender os obstáculos enfrentados ao avaliar a aprendizagem devido ao aumento do uso de ferramentas de Inteligência Artificial Generativa na criação de textos. As investigações revelaram que, embora haja tecnologias promissoras para identificar IA, nenhuma delas demonstrou resultados completamente consistentes ou confiáveis em todas as situações. A diferença nas porcentagens e as várias interpretações sugerem que essas ferramentas ainda estão em processo de aprimoramento e devem ser empregadas cuidadosamente.

O presente estudo não teve a intenção de depreciar as ferramentas existentes ou sugerir que seu uso seja descontinuado. Pelo contrário, nosso objetivo foi oferecer insights e elementos que apontam para a oportunidade de aprimorar essas tecnologias, e, eventualmente, impulsionar o desenvolvimento de novas soluções capazes de identificar, com maior precisão, conteúdos gerados por Inteligência Artificial.

As ferramentas *ZeroGPT*, *Copyleaks* e *Quillbot* foram mais eficazes na detecção de textos gerados por IA, enquanto outras como *Smodin.io* e *Rephrase* mostraram maior

tolerância, falhando em identificar corretamente textos que eram totalmente artificiais. Essas inconsistências destacam a importância de uma abordagem mista que incorpore a análise humana como um complemento à avaliação automatizada, conforme indicado por Silva e Gomes (2023) e Liang et al. (2023).

Para pesquisas futuras, sugere-se aumentar a amostra de textos, incluir um maior número de ferramentas de detecção, bem como aprofundar-se em abordagens pedagógicas que vejam a presença da IA como uma possível aliada e não apenas como um obstáculo à avaliação tradicional.

Referências

Almeida, J. C. P. (2023). Textos Gerados por Inteligência Artificial e suas Implicações no EAD. EaD em Foco, v. 13, n. 1, e2083.

Cervo, A. L., Bervian, P. A. and Silva, R. H. (2006). Metodologia científica. 6. ed. São Paulo: Pearson Prentice Hall.

Copyleaks. (S.D.). AI Detector. Available at: https://copyleaks.com/ai-content-detector. Accessed on: June 12, 2025.

Detecting-ai. (S.D.). AI Detector & AI Checker for ChatGPT, Gemini, Claude & mais. Available at: https://detecting-ai.com/pt. Accessed on: June 12, 2025.

Gao, C. et al. (2023). TextDetox: Robust Detection of AI-Generated Text via Content Preservation. arXiv preprint. Available at: https://arxiv.org/abs/2305.11676. Accessed on: June 12, 2025.

Justdone. (S.D.). AI Detector Justdone. Available at: https://justdone.com/ai-detector. Accessed on: June 12, 2025.

Liang, P. et al. (2023). Can AI Detectors Tell Who Wrote This? Limitations of Current Tools. arXiv preprint. Available at: https://arxiv.org/abs/2305.13217. Accessed on: June 12, 2025.

Liu, R. et al. (2020). Impacts of the digital transformation on the environment and sustainability. European Commission. Available at: https://op.europa.eu/en/publication-detail/-/publication/5b17f1f5-b558-11ea-bb7a-01aa75ed71a1. Accessed on: June 12, 2025.

Pará, J. (2024). A história do Brasil: períodos e principais eventos. Toda Matéria. Available at: https://www.todamateria.com.br/historia-do-brasil/. Accessed on: June 10, 2025.

Quillbot. (S.D.). Detector de IA. QuillBot. Available at: https://quillbot.com/ai-content-detector. Accessed on: June 12, 2025.

Rephrase. (S.D.). Reescrita/detector de IA. Available at: https://www.rephrase.info/br/detector-de-ia. Accessed on: June 12, 2025.

Silva, M. and Gomes, A. (2023). O uso de detectores de texto gerado por IA na educação: limites e possibilidades. Revista Tecnologias na Educação, v. 9, n. 2.

Smodin.Io. (S.D.). AI Detector. Available at: https://smodin.io/ai-content-detector. Accessed on: June 12, 2025.

Zerogpt. (S.D.). Detector de conteúdo ChatGPT/GPTZero. ZeroGPT. Available at: https://zerogpt.com/. Accessed on: June 12, 2025.